

AI Contract Benchmarks Look Authoritative. But Are They Missing Your Deal?

Commercial and Contracts

Technology, Privacy, and eCommerce



Banner artwork by Shutterstock AI / *Shutterstock.com*

Cheat Sheet

- **Check the data.** Many AI benchmarking tools rely on SEC-filed contracts, which may exclude private deals and the most relevant comparables.
- **Question “market standard.”** Ask vendors how benchmarks are sourced, filtered, and supported before relying on their conclusions.

- **Keep strategy human.** AI can flag clauses and deviations, but negotiation decisions still require legal judgment and deal context.

- **Starting point, not final answer.** AI contract analysis can accelerate review, but it should inform — not replace — counsel's assessment of risk, leverage, and commercial goals.

The pitch is compelling. Upload your contract, run it through an AI tool, and within seconds, you have a benchmark: which clauses deviate from market standard, where you are exposed, and what comparable parties typically agree to. For in-house counsel managing high contract volumes with lean teams, this sounds like exactly the kind of leverage technology should provide.

The legal technology vendors have responded accordingly. The release of Stanford Law School's Material Contracts Corpus (MCC) in April 2025, which assembles over one million contracts structured for AI use, has given the industry a flagship resource and added academic credibility to claims that AI can reliably identify what sophisticated parties agree to. Bloomberg Law's Draft Analyzer expressly markets benchmarking against "market-standard language filed in EDGAR documents," EDGAR being the U.S. Securities and Exchange Commission's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) filing system.

These tools are genuinely useful when applied to appropriate tasks. However, there is a gap between their reliable capabilities and how they are marketed. Before relying on AI-generated benchmarks for negotiation strategy, understand both the contents and limitations of the underlying databases.

> Before relying on AI-generated benchmarks for negotiation strategy, understand both the contents and limitations of the underlying databases.

Where the data comes from

The Stanford MCC and other comparable public contract datasets are drawn from the same source: contracts filed with the U.S. Securities and Exchange Commission by publicly traded companies. Under SEC rules, public companies must disclose contracts that are "material" to their business as exhibits to their periodic filings. That is the universe of available data. Every contract in these databases got there because a public company was required to disclose it.

This is a foundational aspect of the dataset and has significant implications for what the data can and cannot reveal.

What is missing, and why it matters

Private company deals

Contracts between private companies do not appear in these databases. Agreements involving private equity-backed portfolio companies, large private enterprises, and family-owned

conglomerates are absent from the MCC and similar public corpora. Many of the most heavily negotiated commercial agreements occur within the private sector and remain inaccessible to AI benchmarking tools.

The agreements that actually got negotiated between equals

This is the most significant gap. Contracts negotiated between large, sophisticated parties with equal bargaining power are most likely to remain private. These transactions, handled by experienced practitioners, are precisely the comparables that parties need for effective benchmarking.

AI tools benchmark against disclosed contracts, not against current agreements between industry peers. As a result, benchmarks may reflect terms accepted by less powerful parties rather than those negotiated between equals.



The leverage dynamics problem

Collectively, these gaps create a systematic distortion in what AI tools present as “market standard.” The database is skewed toward transactions disclosed due to materiality for smaller or more dependent parties, so benchmarks may reflect leverage dynamics favoring less powerful parties.

AI tools do not disclose these limitations when reporting that an indemnification structure, limitation of liability, or termination provision reflects “market standard.” The tool cannot indicate that its benchmarks may be based on transactions with leverage dynamics that differ significantly from your own.

What do these tools do well

To be direct: this article is not arguing that SEC-filed contract corpora are without value, nor that AI contract tools should be dismissed. They are genuinely valuable for a range of tasks.

For clause discovery and precedent research, public corpora are excellent. If you want to understand how indemnification provisions are typically structured in public technology licensing agreements or how earn-out definitions have evolved over time in M&A transactions, the MCC and comparable datasets are rich and searchable resources. For trend analysis, a large corpus of disclosed contracts is exactly the right tool for understanding how market practice in a particular area has shifted over a decade. For public company comparables, where your counterparty is itself a public company and the disclosed universe is directly on point, AI benchmarking adds genuine value.

> To be direct: this article is not arguing that SEC-filed contract corpora are without value, nor that AI contract tools should be dismissed. They are genuinely valuable for a range of tasks.

The argument here is narrower and more specific: when an AI tool claims to identify a “market standard” for a high-stakes transaction between parties, the underlying database may not contain the comparables that matter. The tool’s confidence can exceed the reliability of its data. That gap is what in-house counsel needs to understand: the distance between what the tool claims and what the data supports.

Practical guidance for In-house counsel

Calibrate your reliance to the stakes and the counterparty

AI is best suited for standardized agreements where the relevant market is well represented in public filings. Look for tools that benchmark against your internal playbook or curated clause standards. For major transactions involving significant acquisitions, joint ventures, major licensing deals, or material agreements with large, sophisticated counterparties, use AI benchmarks as a research starting point rather than a negotiation anchor. The most relevant comparable transactions are likely not included in the database.

> Look for tools that benchmark against your internal playbook or curated clause standards.

Ask where does the tool’s data come from

Before relying on an AI benchmarking claim, ask the vendor directly and specifically:

- What percentage of the benchmark set comes from EDGAR or other SEC-filed sources?
- Can results be filtered by company size, industry, governing law, deal type, transaction date, and party posture?
- Does the tool distinguish between internal playbook compliance and external market benchmarking?
- Does it show the source documents underlying each benchmark claim, so you can evaluate the comparables yourself?
- Is any private contract data included in the corpus, and if so, how was it obtained and from what population?

Vendors who cannot answer these questions clearly, or who promote “over one million contracts” without disclosing that all are SEC filings, are not providing the information necessary to evaluate their product.

Preserve the strategic layer

AI tools should manage the mechanical aspects of contract practice, such as organizing documents, flagging deviations from your playbook, identifying missing provisions, and comparing drafts. Experienced practitioners must retain responsibility for the strategic layer, including deciding which deviations to address, sequencing negotiations, determining when to concede, and interpreting the counterparty. AI can provide benchmarks, but the value of experienced counsel lies in applying that information to the specific parties, dynamics, and stakes of each transaction.

Be skeptical of “market standard” claims in fast-moving areas

AI tools are pattern-matchers trained on historical data. In fast-moving or novel areas, the most current market practice is being negotiated in private right now and will not appear in any public corpus for years, if ever. This applies with particular force to AI licensing, data rights, complex technology transactions, and new regulatory frameworks. In these areas, especially, “market standard” from an AI tool may reflect practice that is already outdated.

The bottom line

AI contract benchmarking is a powerful research aid. But when it labels a term “market standard,” in-house counsel should ask whether that standard reflects the deals most comparable to theirs, or only the contracts that public companies were required to disclose. That is not a reason to stop using these tools. It is a reason to use them with the same critical judgment you would apply to any data source whose methodology you have not yet examined.

The best in-house practitioners will use AI tools for what they are genuinely good at; they will not mistake document analysis for the strategic judgment that no corpus of contracts can replace.

Join ACC for more AI insights!

Disclaimer: The information in any resource in this website should not be construed as legal advice or as a legal opinion on specific facts, and should not be considered representing the views of its authors, its authors’ employers, its sponsors, and/or ACC. These resources are not intended as a definitive statement on the subject addressed. Rather, they are intended to serve as a tool providing practical guidance and references for the busy in-house practitioner and other readers.

Gregory Kline



Senior Director, IP Counsel

Thermo Fisher Scientific

Gregory Kline is Senior Director, Intellectual Property and Transactions Counsel at Thermo Fisher Scientific, a Fortune 500 company headquartered in Waltham, Massachusetts. He previously practiced at Morgan, Lewis & Bockius LLP before moving in-house, and now advises on intellectual property, licensing, and complex transactions in the biotechnology and life sciences sectors. He earned a bachelor’s degree from the University of California, San Diego, a Ph.D. in Biology/Immunology from the University of California, Los Angeles, and a juris doctor from the University of California College of the Law, San Francisco (formerly UC Hastings). He serves as a volunteer legal officer with the Civil Air Patrol, supporting its mission of developing youth leadership and mentoring future leaders, and as an election officer in San Francisco.